

Extinction and renewal of Pavlovian modulation in human sequential Feature Positive discrimination learning

Frank Baeyens,¹ Debora Vansteenwegen, Tom Beckers, Dirk Hermans, Ineke Kerkhof, and Annick De Ceulaer

Department of Psychology, University of Leuven, B-3000 Leuven, Belgium

Using a conditioned suppression task, we investigated extinction and renewal of Pavlovian modulation in human sequential Feature Positive (FP) discrimination learning. In Experiment 1, in context *a* participants were first trained on two FP discriminations, $X \rightarrow A+/A-$ and $Y \rightarrow B+/B-$. Extinction treatment was administered in the acquisition context *a* (*aaa* group) or in a new context *b* (*aba* group), and comprised $X \rightarrow A-$ extinction and $Y-$ control trials. Discriminative $X \rightarrow A/A$ responding was lost in both groups when tested in the extinction context, but partially recovered in the *aba* and not in the *aaa* group when tested in the acquisition context, suggesting extinction and renewal of extinguished modulation. The same was observed for the $Y \rightarrow B/B$ control pair, however, questioning whether the loss of discriminative $X \rightarrow A/A$ responding represented genuine extinction of modulation. In Experiment 2, including only *aba* groups, participants were trained in context *a* on two FP discriminations, $X \rightarrow A+/A-$ and $Y \rightarrow B+/B-$, after which the group "Extinction" was exposed to $X \rightarrow A-$ extinction trials in context *b*, whereas the group "Control" was exposed to $X-$ control trials; concurrently, both groups received further $Y \rightarrow B+/B-$ training. In the group Control, differential $Y \rightarrow B/B$ and $X \rightarrow A/A$ responding were acquired and maintained throughout the experiment. In the group Extinction, while $Y \rightarrow B/B$ responding was also maintained throughout, differential $X \rightarrow A/A$ responding disappeared because of $X \rightarrow A-$ extinction treatment when tested in the extinction context *b*, but partially reappeared when tested in the acquisition context *a*. This evidences *aba*-renewal of extinguished modulation.

In a Pavlovian sequential Feature Positive (FP) discrimination task, a conditioned stimulus A (the "target" stimulus) is followed by presentation of an unconditioned stimulus (US) only if target A is preceded by another stimulus X (the "feature" stimulus), hence $X \rightarrow A+/A-$. In animal studies, this conditioning schedule has been found to result in feature X becoming a "facilitator" or "positive occasion setter" that controls the behavioral expression of the association between the target A and the US (Holland 1992). Lamoureux et al. (1998) (see also Holland 1992) argued that such a modulatory strategy is invoked in sequential FP training because the stimulus that is crucial for an adequate resolution of the discrimination, that is, feature X, does not succeed in acquiring a direct excitatory association with the US. That is, the less valid US predictor A will, because of its relative advantage over X in terms of temporal contiguity with the US, overshadow the opportunity for the more valid cue X (or its memory trace) to acquire the required direct excitatory associative strength. In order to resolve this impasse, a strategy is deployed whereby an excitatory A-US association is formed that is modulated by feature X.

Two special characteristics of occasion setters distinguish them from simple Pavlovian excitors and theoretically justify treating them as a functionally different class of Pavlovian stimuli. First, the modulatory powers of a feature X cannot be reduced to, and are in principle independent of, X's own direct associations with the US. This we may call the orthogonality relation between the simple excitatory and the modulatory pow-

ers of a stimulus. If FP discriminations result in occasion setting, this orthogonality is evidenced by the fact that a feature extinction procedure ($X-$), which (behaviorally) abolishes X's own direct excitatory powers, leaves the modulatory powers of X intact (Holland 1983, 1989a, 1991; Rescorla 1986).

A second special characteristic of modulators shows up when assessing their effect on targets (B) different from the one with which they were originally trained (A). Simple Pavlovian excitors linearly and additively combine their influence with the excitatory or inhibitory value of other CSs. For example, an excitator X enhances conditioned responding on XB compound trials compared to the response evoked by B alone, irrespective of the acquisition history of B. Quite differently, the transfer potential of modulators is more limited and selective. Apparently, B targets are only susceptible to a modulator's influence, to the extent that they have become ambiguous because of their own acquisition history (Swartzentruber 1995) and hence are embedded in a representational structure that supports the mode of action of modulators. This ambiguity may result from B having been reinforced and then (partially) extinguished (e.g., Rescorla 1985, 1987, 1991; Swartzentruber and Rescorla 1994; but for counter-evidence, see, e.g., Holland 1986a; Lamarre and Holland 1987); transfer appears to be most consistently produced, however, when B has been the target of modulation by another feature Y ($Y \rightarrow B+/B-$ training) (e.g., Rescorla 1985; Holland 1986b, 1989a,b; Lamarre and Holland 1987).

Different from simple Pavlovian excitors, occasion setters are commonly conceived of as acting on the target A-US association rather than directly on the representation of the US (for reviews, see Holland 1992; Swartzentruber 1995). As proposed in the connectionist model of Pavlovian modulation developed by

¹Corresponding author.

E-mail Frank.Baeyens@psy.kuleuven.ac.be; fax 32 16 3260 99.

Article and publication are at <http://www.learnmem.org/cgi/doi/10.1101/lm.89905>.

Lamoureux et al. (1998) and Schmajuk et al. (1998) (see also Bouton and Nelson 1998, for a very similar model), a simplified depiction of the representational structure presumed to underlie the resolution of sequential FP discriminations is shown in Figure 1A. The target stimulus A acquires both a direct excitatory and a gated inhibitory association with the US; the gated inhibitory association is composed of an excitatory association between the target A and an intermediate control element (Estes 1976) or a hidden unit (Schmajuk et al. 1998), and an inhibitory association between this hidden unit and the US representation. Of crucial importance, the gated inhibitory association can be inhibited by input from feature X to the control element. In this view, responding to the target is ordinarily inhibited by its own gated inhibitory association; the effect of the feature is to cancel that inhibition. It should be noted that this model of sequential FP learning does not preclude that the feature also enters into a direct excitatory association with the US (not shown in the figure); it only stresses that such a direct excitatory feature-US association is not the main mechanism by which the feature controls behavior on $X \rightarrow A$ trials, and that such a direct association is in principle independent of the feature's inhibitory association with the inhibitory target-US association. Because the feature stimulus operates on the inhibitory target-US association rather than on the US representation directly, this model of sequential FP learning is consistent with the major findings on positive occasion setting. First, the model correctly predicts responding to

the target A only if it is preceded by feature X, while feature X by itself should evoke little or no conditioned responses. Second, it explains why the removal of any excitatory strength accrued to X should not affect behavior on the sequential $X \rightarrow A$ compound: feature extinction removes the influence of the direct feature-US association, but leaves the inhibitory feature-hidden unit association intact. Third and finally, it also addresses the selective transfer issue in that it makes clear that new targets B are only susceptible to the feature's influence to the extent that they are embedded in a representational structure that supports the mode of action of positive occasion setters: A new target B should have a gated inhibitory association with the US that the feature X can inhibit.

Whereas a lot of research has been devoted to an analysis of the conditions favoring the acquisition of facilitation, and of the functional characteristics of positive occasion setters once established, comparatively little research has explicitly focused on the conditions and the nature of the processes involved in the extinction of positive modulation, that is, in the reduction or abolishment of the feature's previously acquired capacity to augment responding to the target stimulus. In one of the few systematic analyses of the conditions that remove positive modulation, Rescorla (1986) used an autoshaping preparation with pigeons and first trained the animals on a sequential FP schedule ($X \rightarrow A + / A -$). In line with what we discussed above, he observed that subsequent nonreinforcement of the feature alone ($X -$ treatment) did not affect its modulatory ability (Rescorla 1986, Exp. 2). A reversal of the acquisition contingencies ($A + / X \rightarrow A -$ treatment) (Rescorla 1986, Exp. 1), or nonreinforced presentations of the sequential feature-target compound without separate reinforcement of the target ($X \rightarrow A -$ treatment) (Rescorla 1986, Exp. 3), did markedly attenuate the feature's modulatory power, however. A similar loss of the feature's modulatory ability was observed (Rescorla 1986, Exp. 4) if the feature X was nonreinforced, not in the presence of the original target A, but in sequential compound with a separately trained excitator ($B +$) ($X \rightarrow B -$ treatment); if X was nonreinforced in the presence of a separately trained CS- ($B -$) or a target B that had not been presented before (B^0), the feature's modulatory power remained unaffected, however. Finally, making the feature irrelevant to reinforcement by reinforcing both the target alone and the target preceded by the feature ($X \rightarrow A + / A +$ treatment) was observed not to affect the feature's occasion-setting properties (Rescorla 1986, Exp. 5). Consequently, this series of experiments suggests that it is the conjunction of nonreinforcement and an excitatory target that reduces a feature's positive occasion-setting properties.

At a behavioral level, nonreinforced presentations of the sequential feature-target compound reduce a feature's facilitatory powers. However, this observation does not inform us yet about the exact reason for the loss of X's capability to augment responding to the target A. A first, simple and straightforward possibility might be that the loss of X's facilitatory power is based on associative strength revision, and hence reflects genuine unlearning of the previously acquired modulatory influence on the target A-US association. For example, in terms of the model of positive occasion setting outlined above (Fig. 1A), this would imply that the inhibitory association between the feature and the hidden unit weakens or disappears, such that the (gated) inhibitory target-US association regains control and conditioned responding disappears. The alternative possibility is that, rather than reflecting unlearning, the loss of the feature's modulatory power is a performance phenomenon: Perhaps the associative structure underlying the feature's facilitatory power remains largely intact, but is supplemented with new, corrective associations that counteract the effect and the (behavioral) expression of the original associative structure.

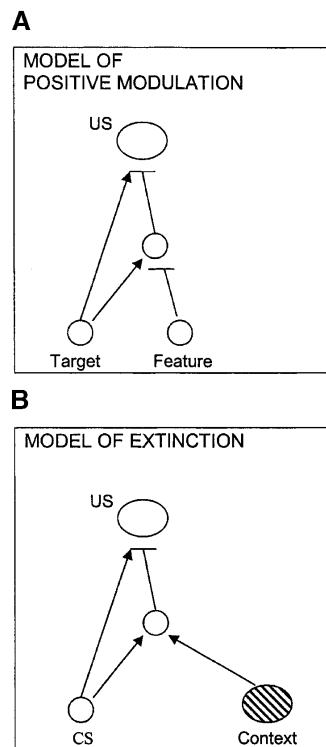


Figure 1. (A) Model of positive modulation. The Target stimulus acquires a direct excitatory association (arrow) and a gated inhibitory association with the US; the gated inhibitory association is composed of an excitatory association between the target and an intermediate hidden unit, and an inhibitory association between this hidden unit and the US (blocked line). The gated inhibitory association can be inhibited by input from the Feature stimulus to the hidden unit. (B) Model of extinction. During acquisition, the target CS acquires a direct excitatory association with the US. During extinction a gated inhibitory association is established between the CS and the US; the intermediate hidden unit requires input from the CS and from the extinction Context for activation of the inhibitory association between the control element and the US.

The latter hypothesis is no more than a logical extension of what is known about extinction of simple Pavlovian exciters, namely, current theories generally assume that during extinction, rather than unlearning the previously acquired excitatory CS-US association, a new inhibitory association is formed to the CS (e.g., Wagner 1981; Pearce 1987). Moreover, as can be shown in renewal experiments, there is strong evidence that this new inhibitory association is highly context-specific. "Renewal" refers to the observation that when a CS that was completely extinguished in a context *b* different from the acquisition context *a* is reintroduced in the acquisition context *a*, a recovery or renewal of conditioned responding toward the CS can be observed (e.g., Bouton and King 1983; Bouton 2000, 2002). The renewal phenomenon has been documented both in aversive and in appetitive conditioning preparations (Bouton and Peck 1989), in situations presenting up to 84 extinction trials after only eight acquisition trials (Bouton and Swartzentruber 1989) and using either the physical environment (room cues such as odors or tactile or visual characteristics of the cages) or an internal physiological state (e.g., the effect of alcohol or tranquilizer) as context shifts (Cunningham 1979; Bouton et al. 1990). At a process level, it has been demonstrated that renewal does not depend on the presence of residual excitatory associative strength between the test context *a* and the US (Bouton and King 1983; Bouton and Swartzentruber 1986): Extensive exposure (extinction procedure) to the test context *a* does not affect renewal, and renewal can also be observed when the CS is tested in a new neutral context *c* after extinction in context *b*. Neither is there any evidence for the hypothesis that renewal should depend on the extinction context *b* acquiring inhibitory associative strength and protecting the CS from losing its excitatory strength during the extinction phase (Bouton and King 1983). Rather, the renewal phenomenon suggests that Pavlovian extinction results from the acquisition of a context-specific inhibitory target CS-US association, countering the effect of the excitatory CS-US association. In consequence, extinction performance only (fully) shows up when the CS is tested in the context in which the new inhibitory association was learned. A plausible (but simplified) model of what is learned in extinction of Pavlovian excitation may, therefore, be the representational structure depicted in Figure 1B (Bouton 1993). During acquisition, the target CS *A* acquires a direct excitatory association with the US. During extinction, an additional, gated inhibitory association is established between the target CS and the US; the intermediate control element or hidden unit requires input from both the target CS and the extinction context (Cx^{EXT}) for activation of the inhibitory association between the control element and the US. Hence, when the extinguished CS is tested outside the extinction context, the inhibitory link is not activated, and "renewed" responding occurs. In this way, the context behaves like a (negative) modulator of the target *A*-US association, and extinction of Pavlovian excitation can be conceptualized as an instance of contextual (negative) occasion setting, $A+/Cx^{EXT}(A-)$.

As was suggested above, the question then arises whether something similar may occur when the conceptual equivalent of extinction is applied to a positive occasion setter. Thus, when after sequential FP training the feature is next presented contingently with nonreinforced target presentations ($X \rightarrow A-$), the observed loss of modulatory power of *X* similarly may be specific to the context of "extinction-of-modulation," rather than to imply a permanent erasure of the modulatory ability. Hence, the hypothesis advanced here is that, just as behavioral extinction of excitation reflects the operation of a contextually controlled new association that opposes the effect of the original association, extinction-of-modulation may be established representationally by the creation of a kind of higher-order extinction structure,

consisting of a contextually controlled new association opposing the effect of the original modulatory association between the feature and the target-US link. If this were true, extinction-of-modulation would boil down to an instance of contextual modulation-of-modulation. The plausibility of this hypothesis is increased by the fact that, given concurrent context-dependent (Cx^1/Cx^2) FP reversal training type [$Cx^1(X \rightarrow A+/A-)$, $Cx^2(X \rightarrow A-/A+)$], modulation-by-context of the modulatory ability of *X* has, indeed, successfully been demonstrated (e.g., Nakajima 1994, 1998), suggesting that the effect of $X \rightarrow A-$ presentations can—at least in principle—be controlled by the context.

The purpose of the research reported in this paper was twofold. First, we wanted to replicate in a human Pavlovian-conditioning preparation that nonreinforced feature-target presentations ($X \rightarrow A-$), but not simple nonreinforced presentations of the feature alone ($X-$), result in a loss of a positive occasion setter's modulatory power. Second, we wanted to investigate whether this (behavioral) loss of modulatory power reflects unlearning or contextually controlled additional learning opposing the effect of the original modulatory association. A human Pavlovian-conditioning preparation that appears to be well-suited to investigate these issues was recently described by Baeyens et al. (2001, Exp. 1, 2, and 2A). Using an online conditioned suppression preparation (the "Martians" computer game developed by Arcediano et al. 1996), Baeyens et al. (2001) created the conditions that animal research has identified as optimal for inducing positive occasion setting in FP discrimination learning (sequential presentation of feature and target stimuli of different sensory modalities, with an empty time gap between both stimuli) (see Holland 1992; Swartzentruber 1995), and, indeed, observed a pattern of responding that is typical for modulation. Participants showed strong conditioned responding to the target *A* only if it was preceded by feature *X*, while feature *X* by itself evoked little or no responding. Also, the serially trained Positive Feature showed a strong transfer selectivity when combined with new targets *B*, in that only a target *B* that had been involved in another sequential FP discrimination was affected by feature *X*, but not a neutral, a partially reinforced, or a conditioned and then extinguished target *B*. Finally, and of direct relevance to our present concerns, the feature *X*'s modulatory power was unaffected by a subsequent extinction manipulation ($X-$ presentations), evidencing the orthogonality relation between the excitatory and the modulatory potential of the positive occasion setter. In the two studies reported here, the same conditioning preparation was used to study acquisition, extinction, and renewal of positive modulation.

Experiment 1

The top half of Table 1 summarizes the design of Experiment 1, which on a between-subject base included an *aba* renewal and an *aaa* control group. Both groups were first concurrently trained in context *a* on two sequential FP discriminations, $X \rightarrow A+/A-$ and $Y \rightarrow B+/B-$, such that targets *A* and *B* presented alone were never followed by the US, whereas the sequential $X \rightarrow A$ or $Y \rightarrow B$ compounds were reinforced on 80% of the trials. The feature (*X* or *Y*) and the target (*A* or *B*) were neutral stimuli from different sensory modalities (visual/auditory), and on sequential compound trials ($X \rightarrow A$ or $Y \rightarrow B$) there was an empty time gap between feature and target. The subsequent extinction treatment consisted of nonreinforced $X \rightarrow A-$ extinction trials and of nonreinforced $Y-$ control trials (within-subject), which were presented either in the acquisition context *a* (*aaa* group) or in a new context *b* (*aba* group). Similar to what was done in a study by Matute and Pineno (1998), contexts *a* and *b* were created by manipulating the way in which the "Martians" and "explosions" of the game

Table 1. Experimental design**Experiment 1**

Group	Acquisition Cx <i>a</i>	Acquisition Test Cx <i>a</i>	Acquisition' Cx <i>a</i>	Transfer Acquisition Test Cx <i>a/b</i>	Extinction Cx <i>a/b</i>	Extinction Test Cx <i>a/b</i>	Transfer Extinction Test Cx <i>a</i>
<i>aaa</i>	12 X→A+, 3 X→A−, 15 A−,		4 X→A+, 1 X→A−, 5 A−,	In Cx <i>a</i> X→A−, A−, X−, Y→B−, B−, Y− (2 blocks)	In Cx <i>a</i> 20 X→A− 20 Y−	In Cx <i>a</i> X→A−, A−, X−, Y→B−, B−, Y−, X→B−, Y→A−	X→A−, A−, X−, Y→B−, B−, Y−, X→B−, Y→A−
<i>aba</i>	12 Y→B+, 3 Y→B−, 15 B−	X→A−, A−, X−, Y→B−, B−, Y−	4 Y→B+, 1 Y→B−, 5 B−	In Cx <i>b</i> X→A−, A−, X−, Y→B−, B−, Y− (2 blocks)	In Cx <i>b</i> 20 X→A− 20 Y−	In Cx <i>b</i> X→A−, A−, X−, Y→B−, B−, Y−, X→B−, Y→A−	

Experiment 2

Group	Acquisition Cx <i>a</i>	Acquisition Test Cx <i>a</i>	Acquisition' Cx <i>a</i>	Transfer Acquisition Test Cx <i>b</i>	Extinction Cx <i>b</i>	Extinction Test Cx <i>b</i>	Transfer Extinction Test Cx <i>a</i>
Control	12 X→A+, 3 X→A−, 15 A−,		4 X→A+, 1 X→A−, 5 A−,	X→A−, A−, X−, Y→B−, B−, Y− (2 blocks)	20 X 8 Y→B+, 8 B−	X→A−, A−, X−, Y→B−, B−, Y−, X→B−, Y→A−	X→A−, A−, X−, Y→B−, B−, Y−, X→B−, Y→A−
Extinction	12 Y→B+, 3 Y→B−, 15 B−	X→A−, A−, X−, Y→B−, B−, Y−	4 Y→B+, 1 Y→B−, 5 B−		20 X→A−, 8 Y→B+, 8 B−		

For half of the participants in each group, stimuli X and Y are (counterbalanced) a 1.5-sec presentation of a complex sound pattern1 and a 1.5-sec presentation of another complex sound pattern2, whereas stimulus A and B are (counterbalanced) a 1.5-sec full-screen presentation of a marble green background pattern and a 1.5-sec presentation of a brown-golden background pattern. For the other half of the participants in each group, the visual stimuli functioned as features X and Y, whereas the auditory stimuli were the targets A and B. An "→" sign denotes a sequential presentation of the feature and the target stimuli (interstimulus interval 1.5 sec. A "+" sign represents a reinforced trial (0.5-sec white flashing screen plus complex sound pattern 3), and a "−" sign represents an unreinforced trial. At Test moments, the critical test stimuli are presented for 3 sec rather than for 1.5 sec. Cx *a* and Cx *b* were the two different contexts in which the Martians appeared (vertical or horizontal, counterbalanced).

appeared on the computer screen (horizontally/vertically; see Materials and Methods for more details). In a previous renewal study at our lab, the same operationalization of contexts already proved to support simple *aba*-renewal of extinguished Pavlovian excitation, suggesting that in this computer game the "horizontal/vertical" manipulation is functionally equivalent to the room/cage manipulations in other human and animal renewal research (Baeyens et al. 2005). In the *aaa* and the *aba* groups alike, feature X was expected to lose its facilitatory power when tested in the extinction context, whereas feature Y was predicted to maintain its modulatory power. Finally, in group *aba* (but not in group *aaa*), a return to the acquisition context *a* was expected to cause a partial recovery of X's occasion-setting powers. In the post-extinction tests we assessed the effect of the feature stimuli (X and Y) on their original targets (A and B, respectively) and in a transfer test, on the target of the other conditional discrimination (B and A, respectively).

The two experiments reported in this paper measured online conditioned suppression as an index of learning of the Pavlovian conditional discriminations. In a prototypical animal online conditioned suppression preparation, subjects first learn to emit a regular pattern of operant responding (e.g., bar pressing reinforced by food); next, they are exposed to a superimposed Pavlovian CS-US contingency (e.g., a neutral tone predicting an electrical shock). Presentation of the US immediately and unconditionally suppresses operant responding, whereas the neutral CS originally does not. When the subject learns, however, that the CS reliably predicts US occurrence, conditioned suppression to the CS is typically observed: operant responding is now also disrupted in the presence of the CS. In the Martians preparation, a similar rationale is followed, the major exception being that an instructed US is used. An instructed US is a stimulus that becomes significant and affects behavior by means of prior explicit verbal

instructions, rather than by its innate biological significance for the subject (for details, see Materials and Methods).

Results**General remarks**

Participants' behavior on the critical test trials was expressed in terms of suppression ratios of the form $a/(a + b)$, where *a* is the number of responses during the critical test stimulus, and *b* is the number of responses in an equal period of time immediately preceding the onset of a given trial. For example, on a 3-sec target-alone (A−) or feature-alone (X−) test trial, *b* equals the number of responses in the 3-sec time slot immediately preceding A− or X−, respectively, whereas on a sequential compound X→A− test trial, *b* for target stimulus A equals the number of responses in the 3-sec time slot immediately preceding X. Because we had clear a priori hypotheses, these suppression ratios were analyzed using planned comparisons. Following Kirk (1995), mean square error (MSE) terms and dfs appropriate for the specific contrasts were used. On several occasions, we used contrasts testing within- or between-group differences between differences (e.g., whether the difference between responding to X→A versus to A is different in context *a* from in context *b*), which is statistically identical to testing interactions between the variables under concern. An α level of 0.05 was used for all planned comparisons. Finally, on some occasions, we also examined the presence or absence of "any" suppression to a stimulus, by determining whether the 0.95 confidence interval of the suppression ratio included 0.50.

Acquisition and Transfer of Acquisition tests

Figure 2 depicts mean suppression ratios at the tests of Acquisition, Transfer of Acquisition, Extinction, and Transfer of Extinction for the targets alone, for the targets preceded by the feature,

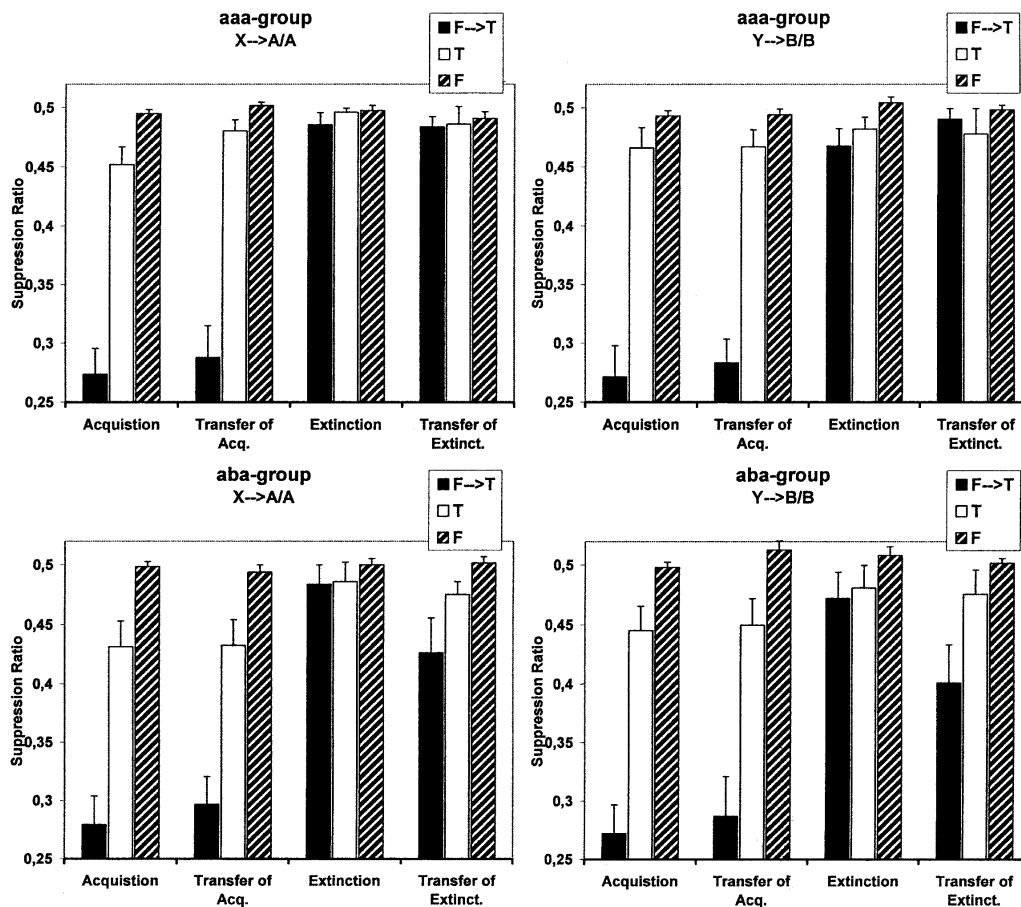


Figure 2. (Experiment 1) Mean suppression ratios at the tests of Acquisition, Transfer of Acquisition, Extinction, and Transfer of Extinction for the targets alone (T), for the targets preceded by the feature (F→T), and for the features alone (F), separately for the X→A/A (left side) and for the Y→B/B (right side) discriminations; the top half of the figure depicts the results of the *aaa* group, the bottom half the results of the *aba* group. Lower values represent stronger conditioned responding. Brackets represent standard errors of the means.

and for the features alone, separately for the X→A/A (left-hand side) and for the Y→B/B (right-hand side) discriminations; the top half of the figure depicts the results of the *aaa* group, the bottom half the results of the *aba* group. In both groups, at the Acquisition test participants reacted with little or no suppression during targets A or B alone, but showed substantial suppression to the same stimuli A or B if they were preceded by the features X or Y, respectively. Features X and Y presented alone did not evoke substantial suppression, and definitely less than on X→A or Y→B sequential compound trials. At the Transfer of Acquisition test, both in the *aaa* and in the *aba* group differential responding to A versus to X→A and to B versus to Y→B seemed to be fully preserved.

Planned comparisons confirmed that in the *aaa* group, the X→A/A discrimination (suppression to target A preceded by feature X vs. suppression to target A presented alone) was reliable at the Acquisition test, $F_{(1,23)} = 34.50$, $MSE = 0.011$, and that it remained so when measured a second time at the Transfer of Acquisition test, $F_{(1,23)} = 36.49$, $MSE = 0.012$. Discriminative X→A/A responding was not reliably different at the two test moments, $F < 1$, $MSE = 0.007$. Similarly, the Y→B/B discrimination was reliable both at the Acquisition test, $F_{(1,23)} = 29.14$, $MSE = 0.016$, and at the Transfer of Acquisition test, $F_{(1,23)} = 45.66$, $MSE = 0.009$, while there was no significant difference in discriminative responding between the two test moments, $F < 1$, $MSE = 0.006$. There was no reliable suppression to features X or Y alone at either test moment.

The same pattern of data was observed in the *aba* group: the X→A/A discrimination was reliable at the Acquisition test in context *a*, $F_{(1,23)} = 25.94$, $MSE = 0.011$, and remained so when measured at the Transfer of Acquisition test in context *b*, $F_{(1,23)} = 35.23$, $MSE = 0.006$. Discriminative X→A/A responding was not reliably different in the two test contexts, $F < 1$, $MSE = 0.006$. Similarly, the Y→B/B discrimination was reliable both at the Acquisition test in context *a*, $F_{(1,23)} = 38.51$, $MSE = 0.009$, and at the Transfer of Acquisition test in context *b*, $F_{(1,23)} = 18.65$, $MSE = 0.017$, while there was no significant difference in discriminative responding between the two test contexts, $F < 1$, $MSE = 0.009$. Finally, there was no reliable suppression to features X or Y alone at any of the test moments.

Extinction and Transfer of Extinction tests: A/X→A/X and B/Y→B/Y trials

At the Extinction test, both in the *aaa* group and in the *aba* group suppression to X→A and discriminative X→A/A responding had completely disappeared (see Fig. 2). In the subsequent Transfer of Extinction test, a difference emerged between both groups: Whereas participants in group *aaa* continued to react with little or no suppression to each of the stimuli, participants in group *aba* showed some renewed suppression to the sequential feature-target compound X→A when tested in context *a*. Unexpectedly, a very similar pattern of results was obtained for the Y→B/B control pair as for the X→A/A pair. Participants in group *aaa* failed to

show any differential responding to B versus to Y→B both at the Extinction and at the Transfer of Extinction tests, whereas participants in group *aba* showed a similar loss of differential Y→B/B responding at the Extinction test in context *b*, followed by a partial return of differential responding at the Transfer of Extinction test in context *a*.

Planned comparisons confirmed that after the nonreinforced X→A – presentations in the *aaa* group, feature X no longer affected responding to target A. Differential X→A/A responding was significantly reduced from the Generalization of Acquisition to the Extinction test, $F_{(1,23)} = 28.02$, $MSE = 0.007$, such that it had completely disappeared at the Extinction test, $F < 1$, $MSE = 0.001$. There was no indication for differential X→A/A responding at the subsequent Transfer of Extinction test, $F < 1$, $MSE = 0.004$, nor for a change in differential X→A/A responding across the final two test moments, $F < 1$, $MSE = 0.002$. Also, there was no reliable suppression to feature X alone at either of the test moments.

Not as expected, planned comparisons also showed that after the nonreinforced Y-alone presentations in the *aaa* group, feature Y no longer affected responding to target B. Differential Y→B/B responding was significantly reduced from the Generalization of Acquisition to the Extinction test, $F_{(1,23)} = 27.50$, $MSE = 0.006$, such that it had also completely disappeared at the Extinction test, $F < 1$, $MSE = 0.003$. There was no indication for differential Y→B/B responding at the subsequent Transfer of Extinction test, $F < 1$, $MSE = 0.007$, nor for a change in differential Y→B/B responding across the final two test moments, $F_{(1,23)} = 1.42$, $p > 0.24$, $MSE = 0.003$. Again, there was no reliable suppression to feature Y alone at either of the test moments.

Similar contrasts in the *aba* renewal group showed that the nonreinforced X→A – presentations resulted in a reliable reduction in differential X→A/A responding from the Generalization of Acquisition to the Extinction test, $F_{(1,23)} = 35.91$, $MSE = 0.003$, such that it had completely disappeared at the Extinction test in context *b*, $F < 1$, $MSE = 0.001$. At the subsequent Transfer of Extinction test in context *a*, however, suppression to X→A had become reliably different from 0.50 again, even though statistically not stronger than to A alone, $F_{(1,23)} = 3.18$, $p < 0.09$, $MSE = 0.009$. Likewise, the increase in suppression to X→A from the Extinction to the Transfer of Extinction test was reliable, $F_{(1,23)} = 4.42$, $MSE = 0.009$, even though the change in differential X→A/A responding across the two test contexts failed to reach a conventional level of significance, $F_{(1,23)} = 2.55$, $p < 0.13$, $MSE = 0.005$. Suppression to feature X alone was not reliably different from 0.50 at either test moment.

Finally, not as expected, a pattern of data indicative for extinction of modulation followed by context-controlled renewal of modulation was also observed for the Y→B/B control pair in group *aba*. The nonreinforced feature Y-alone presentations resulted in a reliable reduction in differential Y→B/B responding from the Generalization of Acquisition to the Extinction test, $F_{(1,23)} = 16.03$, $MSE = 0.009$, such that it had completely disappeared at the Extinction test in context *b*, $F < 1$, $MSE = 0.002$. At the subsequent Transfer of Extinction test in context *a*, however, suppression to Y→B had become reliably different from 0.50 again, and significantly stronger than to B alone, $F_{(1,23)} = 8.17$, $MSE = 0.008$. The increase in suppression to Y→B from the Extinction to the Transfer of Extinction test was reliable, $F_{(1,23)} = 4.75$, $MSE = 0.013$, as was the change in differential Y→B/B responding across the two test contexts, $F_{(1,23)} = 4.28$, $MSE = 0.006$. Suppression to feature Y alone was not reliably different from 0.50 at either test moment.

When the data of the X→A/A extinction and the Y→B/B control pair were aggregated, a final within-group comparison in group *aba* confirmed the reliability of the overall increase in dif-

ferential responding to the target alone (A or B) versus to the target preceded by the feature (X→A or Y→B) when going from the Extinction context *b* to the Transfer of Extinction context *a*, $F_{(1,23)} = 5.37$, $MSE = 0.007$. A between-group comparison finally showed that this interaction between Context (*b/a*) and Stimulus (Target/Feature→Target) was reliably different in the *aba* renewal group from in the *aaa* control group, $F_{(1,46)} = 5.96$, $MSE = 0.005$.

Extinction and Transfer of Extinction tests: X→B/B and Y→A/A trials

Participants' behavior on the two trials assessing the effect of the features on the target of the other conditional discrimination, X→B and Y→A, was compared with the suppression ratios to B and A alone. For feature X, which was expected (and observed) to have lost its modulatory ability because of the preceding X→A – extinction trials, no facilitation of conditioned responding to the transfer target B was expected at the Extinction test; at the Transfer of Extinction test in context *a*, however, renewal of extinguished facilitation eventually might result in the X→B/B discrimination becoming reliable in the *aba* renewal but not in the *aaa* control group. Because feature Y was also observed (even though not expected) to have lost its facilitatory power, similar predictions could be made as for feature X: no differential Y→A/A responding at the Extinction test, and differential Y→A/A responding eventually becoming reliable at the Transfer of Extinction test in the *aba* but not in the *aaa* group.

Figure 3 depicts mean suppression ratios at the tests of Extinction and Transfer of Extinction for the targets alone and for the targets preceded by the feature, aggregated across the X→B/B and the Y→A/A pairs; the top half of the figure depicts the results

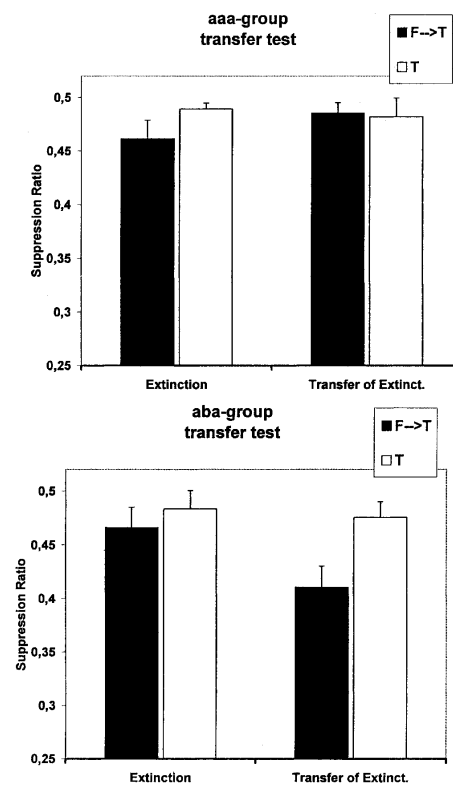


Figure 3. (Experiment 1) Mean suppression ratios at the tests of Extinction and Transfer of Extinction for the targets alone (T) and for the targets preceded by the feature (F→T), aggregated across the X→B/B and the Y→A/A pairs; the top half of the figure depicts the results of the *aaa* group, the bottom half the results of the *aba* group. Lower values represent stronger conditioned responding. Brackets represent standard errors of the means.

of the *aaa* group, the bottom half the results of the *aba* group. Contrast analyses on these suppression ratios confirmed that in the *aaa* group, there was no evidence for differential responding to the sequential feature–target compounds ($X \rightarrow B$ and $Y \rightarrow A$) versus to the targets alone (B and A), neither at the Extinction test, $F_{(1,23)} = 2.45$, $p > 0.13$, $MSE = 0.007$, nor at the Transfer of Extinction test, $F_{(1,23)} < 1$, $MSE = 0.008$. There was no evidence for a change in differential responding across test moments, $F_{(1,23)} = 1.56$, $p > 0.22$, $MSE = 0.007$. In the *aba* group, at the Extinction test responding to the sequential feature–target compounds was likewise not reliably different from responding to the targets alone, $F_{(1,23)} = 3.17$, $p > 0.08$, $MSE = 0.002$. At the subsequent Transfer of Extinction test in context *a*, however, suppression to the sequential feature–target compounds was significantly stronger than to the targets alone, $F_{(1,23)} = 14.33$, $MSE = 0.007$. The increase in suppression to the feature–target compounds from the Extinction to the Transfer of Extinction test was reliable, $F_{(1,23)} = 8.01$, $MSE = 0.009$, as was the change in differential responding across the two test contexts, $F_{(1,23)} = 5.95$, $MSE = 0.005$. In line with the within-group observations in the *aaa* and *aba* groups, a final between-group comparison showed that the latter interaction between Context (*b/a*) and Stimulus (Target/Feature \rightarrow Target) was reliably different in the *aba* renewal group from in the *aaa* control group, $F_{(1,46)} = 6.17$, $MSE = 0.006$.

Discussion

Participants in both groups (*aaa* and *aba*) readily mastered the two conditional discriminations ($X \rightarrow A+/A-$ and $Y \rightarrow B+/B-$): they learned to inhibit bar-pressing on target presentations conditional on the presence of the sequential features, whereas the features by themselves evoked no suppression of the bar-pressing behavior (Acquisition test). This pattern of data suggests that the sequential FP discriminations were resolved by feature-modulated target–US associations rather than by the features becoming simple excitors.

The features' acquired facilitatory ability was not affected by a context switch, in that discriminative $X \rightarrow A/A$ and $Y \rightarrow B/B$ responding remained fully intact in the *aba* group when assessed at the Transfer of Acquisition test in the new context *b*, just as it remained fully intact when assessed a second time in context *a* in the *aaa* control group. This suggests that Pavlovian facilitation, like Pavlovian excitation, readily transfers to contexts different from the original training context (Bouton 2000). Also, this finding reduces the likelihood of an explanation of eventually observed *aba*-renewal in terms of incomplete extinction of the original excitatory/facilitatory associations: If suppression to $X \rightarrow A$ would only partially transfer to the extinction context *b*, post-extinction renewal of conditioned responding to $X \rightarrow A$ might simply result from residual associative strength showing up again in context *a*.

Just as observed in animal preparations (e.g., Rescorla 1986), nonreinforced feature–target presentations ($X \rightarrow A-$) were clearly effective in (behaviorally) extinguishing the feature's (X) positive occasion-setting powers: at the Extinction test, both in group *aaa* and in group *aba* discriminative $X \rightarrow A/A$ responding had completely disappeared. Contrary to expectations, however, a similar loss of facilitatory power occurred to the control feature Y , which had never been presented conjointly with nonreinforcement of its target B . This finding is especially surprising given that, using an identical procedure, Baeyens et al. (2001, Exp. 1) observed that 10 nonreinforced feature-alone presentations following 16 $X \rightarrow A+$, 4 $X \rightarrow A-$, and 20 $A-$ acquisition trials did not at all affect the feature's facilitatory ability. A possible but admittedly speculative explanation for the unexpected loss of feature Y 's facilitatory power might be in terms of an acquired equivalence effect.

Namely, there is considerable evidence that stimuli sharing a common training history become more similar, such that a new stimulus function acquired by one stimulus easily generalizes to the other stimuli of the set of equivalent stimuli (see, e.g., Honey and Hall 1989; Hall 1996; Bonardi 1998). In the present experiment, participants were trained on two similar sequential FP discriminations, $X \rightarrow A+/A-$ and $Y \rightarrow B+/B-$, such that features X and Y acquired similar positive occasion-setting functions and hence may have become functionally equivalent. As a result of this, the subsequent extinction of X 's modulatory power (caused by the nonreinforced $X \rightarrow A-$ presentations) may have generalized to feature Y , such that conditioned responding to the $Y \rightarrow B$ compound disappeared, even though feature Y had never been presented conjointly with nonreinforcement of its target B . It may be worthwhile to mention that this observation is reminiscent of something that Pavlov (1927) reported with respect to extinction of simple excitors, namely, "secondary extinction": if two CSs, A and B , are first associated with a US, and next one of the two CSs (A) is subject to extinction, diminished conditioned responding to the other CS (B), which was never directly subjected to extinction, can also be observed.

Irrespective of the specific cause of the loss of the feature's modulatory powers, the data of the Transfer of Extinction test clearly showed that this loss was not permanent. In a similar way as extinguished Pavlovian excitation may partially reappear when the CS is tested in a context different from the extinction context, the modulatory ability of the features X and Y was partially restored when assessed in a context *a* different from the extinction context *b*. This *aba*-renewal of modulation suggests that extinction of facilitation probably reflects contextually controlled suppression, rather than permanent erasure, of the feature's modulatory ability.

Finally, the renewed modulatory power of features X and Y not only showed up when tested in combination with the targets they had originally been trained with ($X \rightarrow A$ and $Y \rightarrow B$ trials), but also when combined with the targets of the other conditional discrimination ($X \rightarrow B$ and $Y \rightarrow A$ trials). In other words, outside the extinction context the extinguished features behaved like features that have not been extinguished, in the sense that they also facilitated responding to targets that had been subject to modulation by other feature stimuli.

Materials and Methods

Participants

A total of 48 undergraduate psychology students (34 females, 14 males, age 17–22) participated as a partial fulfillment of course requirements. None of them had any previous experience with the Martians preparation, and they were all uninformed as to the purpose of the experiment. An equal number of participants ($n = 24$) was assigned randomly to the *aaa*-control and to the *aba*-renewal groups. Within each group, for half of the participants, features X and Y were auditory stimuli and targets A and B were visual stimuli, whereas the reverse was true for the other half. Participants were tested individually, and test duration was ~45 min.

Apparatus, software, and stimuli

The experiment was run on an IBM-compatible Pentium 4 2.4-GHz 512-Mb RAM multimedia PC with the participants responding on the spacebar of the keyboard. The Martians preparation was implemented into a flexible Windows 95 environment by Baeyens and Clarysse (1998) using Microsoft Visual C++ 5.0. The critical stimuli used in Experiment 1 were the feature stimuli X and Y , the target stimuli A and B , and the US. For half of the participants in each group, stimuli X and Y were (counterbalanced) a 1.5-sec presentation of a complex sound pattern1 (Windows 95 "Windows 95 maximize.wav," played back in con-

tinuous looping by means of Altec Lansing Multimedia ACS90 computer speakers) and a 1.5-sec presentation of another complex sound pattern² (Windows 95 "Sixties menu command.wav," also played back in continuous looping), whereas stimuli A and B were (counterbalanced) a 1.5-sec full-screen presentation of a marble-green background pattern (Windows 95 "marble.bmp," tiled presentation) and a 1.5-sec full-screen presentation of a patchy brown-golden background pattern (Windows 95 "gold.bmp," tiled presentation). For the other half of the participants in each group, the visual stimuli functioned as features X and Y, whereas the auditory stimuli were the targets A and B. The instructed US consisted of the simultaneous presentation of a 0.5-sec white flashing screen (5 flashes at a rate of 10 flashes/sec; flash-time = 50 msec, inter-flash-time = 50 msec) and another 0.5-sec complex sound pattern³ (Windows 95 "In the computer program error.wav," played back in continuous looping).

Contexts *a* and *b* were created by manipulating the way in which the Martians and explosions appeared on the computer screen. In both contexts, Martians and explosions measured 1.76 cm × 1.76 cm and appeared at intervals of 0.25 sec against a black background, with a space of 2 cm in between each Martian or explosion. In the "horizontal" context, however, they appeared one by one in rows on the screen, from left to right, the rows filling the screen from top to bottom, whereas in the "vertical context" they appeared one by one in columns on the screen, from top to bottom, the columns filling the screen from left to right. Additionally, the color of the Martians and the explosions was different in the two contexts. In the horizontal context a Martian had a green head with red eyes, mouth, and antennae against a yellow background, and an explosion was depicted as a red star against a gray background. In the vertical context, the negative of these colors was used, such that a Martian had a purple head with pale blue eyes, mouth, and antennae against a dark blue background, and the explosion appeared as a pale blue star against a gray background. For half of the participants in group *aba*, the "horizontal" context was context *a* and the vertical context *b*; for the other half of the participants, this assignment was reversed. For half of the participants in group *aaa*, the horizontal context was used throughout all phases of the experiment, whereas the vertical context was used for the other half of the participants.

Procedure

Pretraining

The purpose of the pretraining phase was to teach participants to emit a regular pattern of operant responding, that is, to bar-press consistently using the space bar of the computer keyboard. During this phase, neither feature (X/Y) and target (A/B) stimuli, nor USs were presented. The screen displayed "Martians that were trying to invade the Earth," and the task of the participant was to prevent their landing by pressing the space bar of the keyboard ("shooting a laser gun at each of the Martians that would try to land"). The instructions explained that in case of a hit, an explosion sign rather than a Martian would be displayed; we further stressed the importance of emitting a regular pattern of bar presses, by informing the participants that a new Martian would appear about every half-second, that they should shoot immediately before each Martian would try to land, and that they only had one shot per Martian (for a literal transcription of the instructions, see Baeyens et al. 2001). If the participant pressed the space bar before a new Martian was displayed, the explosion rather than the Martian appeared at that position. The objective was to have as few Martians and as many explosions as possible on the screen. Only one bar press per Martian was allowed. If more than one press was recorded (if the participant's key pressing rate exceeded 4/sec, or when she/he held the space bar down), a Martian rather than an explosion was displayed. The screen was filled when 70 Martians or explosions (10 in each of seven rows [horizontal] or seven in each of 10 columns [vertical]; inter-row and inter-column distance, 2 cm) had been displayed. When the screen was filled, it scrolled up, one line at a time (or scrolled left, one column at a time), to make room for new Mar-

tians, so that there were no "breaks" but a continuous progression between screens. The pretraining phase lasted for 25 sec, so that a total of 100 Martians or explosions appeared on the screen. At the end of this pretraining, the number of hits and misses and the percentage of explosions appeared on a cleared screen.

US-only phase

During this phase, the instructional US was introduced by describing and demonstrating what would happen if participants continued key-pressing during the anti-laser shield (a white flashing screen plus sound): their weapon would temporarily become ineffective, and an inescapable invasion of thousands of Martians would be evoked. During this phase, the USs were not signaled by any other stimulus, and they were presented in the same context as the context of the pretraining phase (horizontal for one-half of the participants, vertical for the other half). A total of four USs were scheduled with an average intertrial interval (ITI) of 7.5 sec (min = 5 sec, max = 10 sec). The US consisted of the 0.5-sec simultaneous presentation of the intermittent white flashing screen plus the complex sound pattern; during the US, the Martians kept appearing on the screen in identically the same way as during the ITIs (4/sec). If no key press was registered during the US, nothing else happened, but if a response was recorded, the US was followed by an invasion. An invasion lasted for 5 sec, during which the background kept flashing (10 flashes/sec), a new complex sound pattern⁴ was played (Windows 95 "Robotz-2.wav," played back in continuous looping), and Martians now invaded the screen with a time interval of 0.1 sec; moreover, bar-pressing was ineffective during an invasion (no explosions appeared contingent on bar-pressing). At the first US trial, the experimenter refrained from pressing the space bar before US presentation; at the second trial, he kept firing during the US in order to produce an invasion; at the third US trial, he again refrained from pressing the space bar and demonstrated that there was no danger in starting firing again immediately after the US had disappeared from the screen; and at the fourth trial, he illustrated that it was important to stop firing before a US was actually presented, by showing the impossibility of timely stopping firing as soon as the US appeared.

Acquisition and Acquisition test

In the acquisition phase, participants were concurrently trained on the two critical sequential FP discriminations, $X \rightarrow A+/A-$ and $Y \rightarrow B+/B-$ (see Table 1, top). In both groups this training took place in the acquisition context *a*, and it was superimposed on the operant baseline task. The switching on of the anti-laser shield (US) was now announced by "indicators" ($X \rightarrow A$ or $Y \rightarrow B$), such that if participants detected the $X \rightarrow A$ -US or $Y \rightarrow B$ -US relations, they could also learn to timely suppress responding (before the US was presented). On trials on which indicators A or B were presented alone, the anti-laser shield was not presented, such that suppression of responding was unnecessary. The experimenter introduced the acquisition phase by telling the participants that they probably understood that it would be very beneficial to have a means to predict the connecting of the anti-laser shield, and that they could do so by paying attention to indicators that would tell when the anti-laser screen was about to be connected, and by learning to distinguish between correct and false indicators. In order to introduce the existence of context different from the one participants had been exposed to so far, they were additionally informed that it might occur that the Martians occasionally made use of new combat groups, deploying new strategies of invasion.

The 15 $A-$, 12 $X \rightarrow A+$, and three $X \rightarrow A-$ trials of $X \rightarrow A+/A-$ training and the 15 $B-$, 12 $Y \rightarrow B+$, and three $Y \rightarrow B-$ trials of $Y \rightarrow B+/B-$ training were subdivided into three equivalent sequential sub-blocks. Each sub-block contained four $A-$, three $X \rightarrow A+$, one $X \rightarrow A-$, four $B-$, three $Y \rightarrow B+$, and one $Y \rightarrow B-$ trials; these were followed by a single presentation of $A-$, $X \rightarrow A+$, $B-$, and $Y \rightarrow B+$ assessment trials. Within each sub-block, the order of presentation of the different trial types was semirandomized, with the restriction that no more than two consecutive trials could be of the same type. In each sub-block, a different random-

ization was used, and each participant was subjected to different trial randomizations. The order of presentation of the assessment trials at the end of each sub-block was also random. The duration of the intertrial interval (ITI) was pseudorandom, with a range between 7.5 and 12.5 sec, and a mean of 10 sec. During the ITIs, the background color remained black. Feature and target presentations were 1.5 sec, and on sequential $X \rightarrow A$ and $Y \rightarrow B$ trials, the empty feature–target interval was fixed at 1.5 sec. On reinforced $Y \rightarrow A+$ and $Y \rightarrow B+$ trials, the US immediately followed the offset of target A/B. On assessment trials, the presentation time of the targets A and B ($X \rightarrow A+$, $Y \rightarrow B+$, $A-$, and $B-$ trials) was lengthened from 1.5 to 3 sec. in order to increase the sensitivity of the suppression measure. Namely, it takes some time to refrain from responding, and additionally participants may also learn that bar-pressing is still safe during early presentation of a (target) stimulus, such that suppression is often rather weak at the beginning of target presentation; it becomes much stronger, however, at the moment that the US should normally occur, that is, after the 1.5-sec time slot of a standard presentation of a (target) stimulus has elapsed.

The Acquisition test followed the third acquisition block without any interruption, and contained the critical measurement of participants' behavior on both sequential FP discriminations in the acquisition context *a*. The targets alone ($A-$ and $B-$), the targets preceded by their feature ($X \rightarrow A-$ and $Y \rightarrow B-$), and also the features alone ($X-$ and $Y-$) were given a single nonreinforced test presentation. Stimulus presentation duration was 3 sec for all targets and for the features presented alone. Each participant received a different random presentation order of these test trials. After this, the number of hits and misses and the percentage of explosions of the complete acquisition series appeared on a cleared screen.

Acquisition was completed with some additional $X \rightarrow A+/A-$ and $Y \rightarrow B+/B-$ training before the start of extinction treatment, such that participants could experience that the experimental contingencies remained the same after the short brake that followed the Acquisition test. These additional acquisition trials were introduced by repeating the main rules of the game, "preventing as many Martians as possible from landing, while being very careful of the anti-laser screen and paying attention to the indicators." The additional acquisition training (Acquisition' in Table 1) comprised a single sub-block of ($X \rightarrow A+/A-$) and ($Y \rightarrow B+/B-$) training, which was identical to the previous acquisition sub-blocks and was presented again in the acquisition context *a*.

Transfer of Acquisition test

Because participants in the *aba* group were to be exposed to the extinction treatment in a context different from the acquisition context *a*, we first assessed to what extent the facilitatory powers of the features *X* and *Y* transferred to this new context *b*. Namely, to the extent that the transfer of the acquired modulatory power of the feature to a new context *b* would be incomplete, several mechanisms might be invoked to explain a possible post-extinction renewal of modulation (see Discussion). Therefore, the targets alone ($A-$ and $B-$), the targets preceded by their feature ($X \rightarrow A-$ and $Y \rightarrow B-$), and also the features alone ($X-$ and $Y-$) were given two nonreinforced test presentations in the extinction context *b*. The Transfer of Acquisition test was subdivided into two sequential sub-blocks each containing one presentation of each trial type in a random order; stimulus presentation parameters were identical to those of the Acquisition test. Participants in the *aaa* control group received the same two sets of nonreinforced test presentations in context *a*.

Extinction and Extinction test

In both groups, participants were exposed to 20 nonreinforced $X \rightarrow A-$ trials and to 20 nonreinforced Y -alone trials, the former treatment aiming at removing feature *X*'s facilitatory power, the latter aiming at leaving *Y*'s facilitatory power intact. The 20 $X \rightarrow A-$ and the 20 $Y-$ trials were subdivided into four equivalent sequential sub-blocks. Each sub-block contained four $X \rightarrow A-$ and four $Y-$ trials, followed by a single $X \rightarrow A-$ and $Y-$ assessment trial. Semirandomization of trial types was accom-

plished in the same way as during acquisition, and the ITIs and stimulus presentation times were identical to those used in the previous phases.

The Extinction test followed the fourth extinction block without any interruption, and indexed participants' extinction behavior on both discriminations in the extinction context (context *b* in the *aba* group vs. context *a* in the *aaa* group). This test included a single nonreinforced presentation of the targets alone ($A-$ and $B-$), of the targets preceded by their feature ($X \rightarrow A-$ and $Y \rightarrow B-$), of the features alone ($X-$ and $Y-$), and also of the targets preceded by the feature from the other discrimination ($X \rightarrow B-$ and $Y \rightarrow A-$). The order of presentation of the eight trials was randomized for each participant, and the ITIs and stimulus presentation times were identical to those used in the previous test phases.

Transfer of Extinction test

The Transfer of Extinction test was similar to the Extinction test, the only difference being that all trials were now presented in the acquisition context *a*. Hence, for participants in the *aba* group this final test involved a switch from the extinction context *b* to the acquisition context *a*, whereas for the participants in the *aaa* group, the final test was an exact replication of the previous Extinction test.

Experiment 2

Experiment 1 demonstrated a partial renewal of facilitation caused by a return to the acquisition context, but delivered surprising results concerning the conditions of the preceding extinction of that facilitatory power. Above we hypothesized that the unexpected loss of modulatory ability of the control feature *Y* might be due to an acquired functional equivalence with feature *X*, which was given extinction treatment. One could also argue, however, that the lack of a difference between the effects of the $X \rightarrow A-$ extinction and the $Y-$ control treatment compromises the very idea of genuine associative extinction-of-modulation, and hence also of renewal of extinguished modulation. Namely, given that the control treatment ($Y-$ trials), which on both theoretical and empirical grounds (Baeyens et al. 2001) should be expected not to remove facilitation, did result in a loss of the feature's modulatory power, the possibility exists that also in the $X \rightarrow A-$ extinction condition the loss of *X*'s facilitatory power was due to nonassociative factors (e.g., the operation of a belief or an abstract rule type, "this phase of the experiment is completely safe" or "USs are no longer administered"), rather than being due to a genuine modification of *X*'s associations. For this reason, we deemed it important to try to replicate the findings on extinction and renewal of modulation under conditions that would allow for a more straightforward interpretation of the loss of facilitatory power.

The bottom half of Table 1 summarizes the design of Experiment 2, which basically replicated the *aba* condition of Experiment 1, but included a between-subject manipulation of the nonreinforced feature→target extinction versus the nonreinforced feature-alone control treatment. Thus, using the same stimuli and the same reinforcement schedule as in Experiment 1, two groups (Extinction and Control) were first concurrently trained in context *a* on two sequential FP discriminations, $X \rightarrow A+/A-$ and $Y \rightarrow B+/B-$. In the subsequent extinction phase, which took place in a new context *b*, participants in the Extinction group received nonreinforced $X \rightarrow A-$ trials, whereas the Control group was exposed to nonreinforced $X-$ trials; in both groups, the $X \rightarrow A-$ or $X-$ trials were intermixed with reinforced $Y \rightarrow B+$ and nonreinforced $B-$ trials. In other words, while the $X \rightarrow A+/A-$ discrimination was either extinguished (Extinction group) or not (Control group), the $Y \rightarrow B+/B-$ discrimination was kept intact. This was mainly done to avoid that participants would start to believe that USs did simply no longer occur in the

second half of the experiment, and would therefore lose suppression to any of the stimuli; at the same time, the inclusion of the $Y \rightarrow B+/B-$ pair allowed for an assessment of the effect of features X and Y on the targets of another conditional discrimination, that is to include some feature transfer tests. Hence in the Extinction group, feature X was expected to lose its facilitatory power when tested in the extinction context, whereas feature Y was predicted to maintain its modulatory power; a return to the acquisition context *a* was expected to cause a partial recovery of X's occasion-setting powers, while feature Y was expected to be unaffected by the context switch. In the Control group, both features X and Y were predicted to maintain their facilitatory powers throughout all phases of the experiment.

Results

Acquisition and Transfer of Acquisition tests

Figure 4 depicts mean suppression ratios at the tests of Acquisition, Transfer of Acquisition, Extinction, and Transfer of Extinction for the targets alone, for the targets preceded by the feature, and for the features alone, separately for the $X \rightarrow A/A$ (left-hand side) and for the $Y \rightarrow B/B$ (right-hand side) discriminations; the top half of the figure depicts the results of the Control group, the bottom half the results of the Extinction group. In both groups, at the Acquisition test in context *a* participants reacted with little suppression during targets A or B alone, but showed substantial

suppression to the same stimuli A or B if preceded by the features X or Y, respectively. Features X and Y presented alone did not evoke any substantial suppression, and definitely less than on $X \rightarrow A$ or $Y \rightarrow B$ sequential compound trials. At the Transfer of Acquisition test in context *b*, both in the Extinction and in the Control group differential responding to A versus to $X \rightarrow A$ and to B versus to $Y \rightarrow B$ seemed to be well preserved.

Planned comparisons confirmed that in the Control group, the $X \rightarrow A/A$ discrimination was reliable at the Acquisition test in context *a*, $F_{(1,23)} = 43.53$, $MSE = 0.010$, and that it remained so when measured at the Transfer of Acquisition test in context *b*, $F_{(1,23)} = 23.12$, $MSE = 0.010$. However, discriminative $X \rightarrow A/A$ responding was slightly but reliably less pronounced in the new test context *b* than in the acquisition context *a*, $F_{(1,23)} = 5.85$, $MSE = 0.003$. The $Y \rightarrow B/B$ discrimination was reliable both at the Acquisition test in context *a*, $F_{(1,23)} = 48.02$, $MSE = 0.007$, and at the Transfer of Acquisition test in context *b*, $F_{(1,23)} = 24.07$, $MSE = 0.013$, while there was no significant difference in discriminative responding between the two test contexts, $F < 1$, $MSE = 0.006$. There was no reliable suppression to features X or Y alone at either test moment.

A very similar pattern of data was observed in the Extinction group: the $X \rightarrow A/A$ discrimination was reliable at the Acquisition test in context *a*, $F_{(1,23)} = 59.96$, $MSE = 0.006$, and remained so when measured at the Transfer of Acquisition test in context *b*, $F_{(1,23)} = 23.45$, $MSE = 0.015$. Discriminative $X \rightarrow A/A$ responding

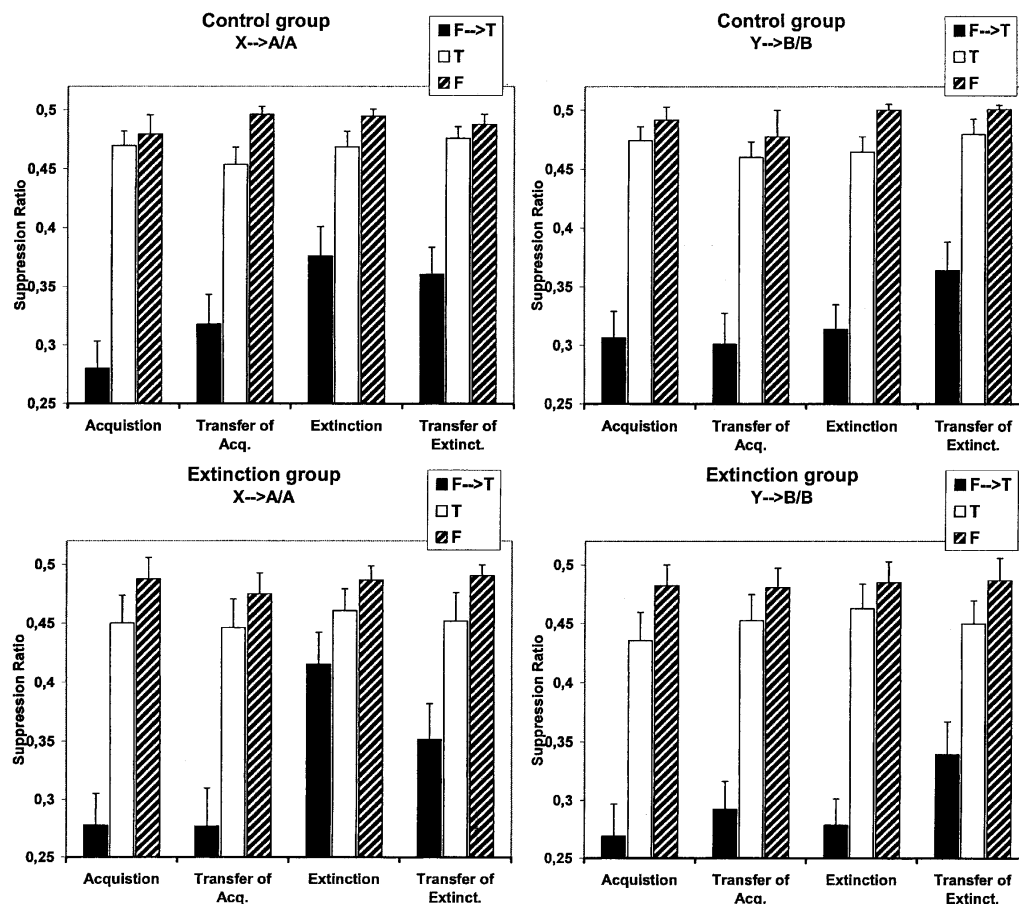


Figure 4. (Experiment 2) Mean suppression ratios at the tests of Acquisition, Transfer of Acquisition, Extinction, and Transfer of Extinction for the targets alone (T), for the targets preceded by the feature ($F \rightarrow T$), and for the features alone (F), separately for the $X \rightarrow A/A$ (left side) and for the $Y \rightarrow B/B$ (right side) discriminations; the top half of the figure depicts the results of the Control group, the bottom half the results of the Extinction group. Lower values represent stronger conditioned responding. Brackets represent standard errors of the means.

was not reliably different in the two test contexts, $F < 1$, $MSE = 0.008$. Similarly, the $Y \rightarrow B/B$ discrimination was reliable both at the Acquisition test in context a , $F_{(1,23)} = 39.38$, $MSE = 0.008$, and at the Transfer of Acquisition test in context b , $F_{(1,23)} = 44.72$, $MSE = 0.007$, while there was no significant difference in discriminative responding between the two test contexts, $F < 1$, $MSE = 0.006$. Finally, there was no reliable suppression to features X or Y alone at either test moment.

Extinction and Transfer of Extinction tests: $A/X \rightarrow A/X$ and $B/Y \rightarrow B/Y$ trials

In the Extinction group, at the Extinction test in context b suppression to $X \rightarrow A$ had become very weak, and discriminative $X \rightarrow A/A$ responding appeared to be seriously reduced (see Fig. 4). However, at the subsequent Transfer of Extinction test in context a , suppression to $X \rightarrow A$ and differential $X \rightarrow A/A$ responding apparently had increased again. At the same time, discriminative $Y/Y \rightarrow B$ appeared to be preserved at both test moments. In the Control group, both discriminative $X \rightarrow A/A$ responding and discriminative $Y \rightarrow B/B$ responding appeared to be intact at the Extinction and at the Transfer of Extinction tests.

Contrast analysis confirmed that in the Control group, the nonreinforced feature alone presentations ($X -$) did not reliably decrease differential $X \rightarrow A/A$ responding from the Generalization of Acquisition to the Extinction test, $F_{(1,23)} = 1.61$, $p > 0.21$, $MSE = 0.007$, such that differential responding to A versus to $X \rightarrow A$ was still reliable at the Extinction test in context b , $F_{(1,23)} = 12.00$, $MSE = 0.009$. At the subsequent Transfer of Extinction test in context a , differential $X \rightarrow A/A$ responding remained reliable, $F_{(1,23)} = 22.24$, $MSE = 0.007$, and it was statistically not different from differential $X \rightarrow A/A$ responding at the preceding Extinction test in context b , $F < 1$, $MSE = 0.007$. Suppression to feature X alone was not reliably different from 0.50 at either test moment.

A similar pattern of results was obtained for the $Y \rightarrow B/B$ pair in the Control group. The continued reinforcement of the sequential feature–target compound ($Y \rightarrow B+$) combined with nonreinforcement of the target alone ($B -$) did not change differential $Y \rightarrow B/B$ responding from the Generalization of Acquisition to the Extinction test, $F < 1$, $MSE = 0.004$, such that differential responding to B versus to $Y \rightarrow B$ remained highly reliable at the Extinction test in context b , $F_{(1,23)} = 44.66$, $MSE = 0.006$. At the subsequent Transfer of Extinction test in context a , differential $Y \rightarrow B/B$ responding was still reliable, $F_{(1,23)} = 13.51$, $MSE = 0.012$, and statistically not different from differential $Y \rightarrow B/B$ responding at the preceding Extinction test in context b , $F < 1$, $MSE = 0.008$. Suppression to feature Y alone was not reliably different from 0.50 at either test moment.

Similar contrasts in the Extinction group showed that the nonreinforced $X \rightarrow A -$ presentations resulted in a reliable reduction in differential $X \rightarrow A/A$ responding from the Generalization of Acquisition to the Extinction test, $F_{(1,23)} = 8.65$, $MSE = 0.011$, such that it was no longer statistically reliable at the Extinction test in context b , $F_{(1,23)} = 4.15$, $p > 0.05$, $MSE = 0.006$. At the subsequent Transfer of Extinction test in context a , however, suppression to $X \rightarrow A$ again had become reliably stronger than to A alone, $F_{(1,23)} = 10.57$, $MSE = 0.011$. Also, the increase in suppression to $X \rightarrow A$ from the Extinction to the Transfer of Extinction test was reliable, $F_{(1,23)} = 5.15$, $MSE = 0.010$, while the increase in differential $X \rightarrow A/A$ responding across the two test contexts failed to reach a conventional level of significance, $F_{(1,23)} = 2.99$, $p < 0.10$, $MSE = 0.006$. Suppression to feature X alone was not reliably different from 0.50 at either test moment.

Analysis of the responses to the $Y \rightarrow B/B$ pair in the Extinction group showed the following. As in the Control group, the

continued reinforcement of the sequential feature–target compound ($Y \rightarrow B+$) combined with nonreinforcement of the target alone ($B -$) did not change differential $Y \rightarrow B/B$ responding from the Generalization of Acquisition to the Extinction test, $F < 1$, $MSE = 0.004$, such that differential responding to B versus to $Y \rightarrow B$ remained highly reliable at the Extinction test in context b , $F_{(1,23)} = 87.79$, $MSE = 0.005$. At the subsequent Transfer of Extinction test in context a , differential $Y \rightarrow B/B$ responding was still reliable, $F_{(1,23)} = 21.76$, $MSE = 0.007$, even though it had become less pronounced than differential $Y \rightarrow B/B$ responding at the preceding Extinction test in context b , $F_{(1,23)} = 7.12$, $MSE = 0.005$. Suppression to feature Y alone was not reliably different from 0.50 at either test moment.

A final within-group comparison in group Extinction confirmed that the interaction between Context (b/a) and Stimulus (Target/Feature \rightarrow Target) was reliably different for the $X \rightarrow A/A$ (extinguished) pair than for the $Y \rightarrow B/B$ (control) pair, $F_{(1,46)} = 6.45$, $MSE = 0.008$.

Extinction and Transfer of Extinction tests: $X \rightarrow B/B$ and $Y \rightarrow A/A$ trials

As in Experiment 1, participants' behavior on the two trials assessing the effect of the features on the target of the other conditional discrimination, $X \rightarrow B$ and $Y \rightarrow A$, was compared with the suppression ratios to B and A alone. In the Control group, both features X and Y were expected (and observed) to have retained their modulatory ability, and therefore facilitation of conditioned responding to the transfer targets was predicted at the Extinction test in context b and at the Transfer of Extinction test in context a . In the Extinction group, feature X was expected (and observed) to have lost its modulatory ability because of the preceding $X \rightarrow A -$ extinction trials, and therefore no facilitation of conditioned responding to the transfer target B was expected at the Extinction test in context b ; at the Transfer of Extinction test in context a , however, renewal of extinguished facilitation eventually might result in the $X \rightarrow B/B$ discrimination becoming reliable. Finally, depending on whether the $X \rightarrow A -$ extinction trials had affected only the modulatory ability of feature X , or also the target A –US associations, participants in the Extinction group might either suppress responding on the $Y \rightarrow A$ trial or not at the Extinction test in context b ; correspondingly, at the Transfer of Extinction test in context a , differential $Y \rightarrow A/A$ responding might either remain or become reliable.

Figure 5 depicts mean suppression ratios at the tests of Extinction and Transfer of Extinction for the targets alone and for the targets preceded by the feature, separately for the $X \rightarrow B/B$ (left-hand side) and for the $Y \rightarrow A/A$ (right-hand side) discriminations; the top half of the figure depicts the results of the Control group, the bottom half the results of the Extinction group. Contrast analyses on the suppression ratios in the Control condition showed that there was evidence for differential responding to B preceded by X versus to B alone both at the Extinction test in context b , $F_{(1,23)} = 10.45$, $MSE = 0.007$, and at the Transfer of Extinction test in context a , $F_{(1,23)} = 15.16$, $MSE = 0.010$. There was no evidence for a change in differential responding across the two test contexts, $F_{(1,23)} = 1.30$, $p > 0.26$, $MSE = 0.006$. Likewise, there was evidence for differential responding to A preceded by Y versus to A alone both at the Extinction test in context b , $F_{(1,23)} = 16.17$, $MSE = 0.013$, and at the Transfer of Extinction test in context a , $F_{(1,23)} = 11.47$, $MSE = 0.007$. There was no evidence for a change in differential responding across the two test contexts, $F_{(1,23)} = 2.09$, $p > 0.16$, $MSE = 0.007$.

In departure from what we expected, a very similar pattern of results was obtained on the $X \rightarrow B/B$ trial in the Extinction group: suppression to B preceded by X was reliably stronger than to B alone at the Extinction test, $F_{(1,23)} = 14.62$, $MSE = 0.004$, and at the Transfer of Extinction test, $F_{(1,23)} = 14.69$, $MSE = 0.007$.

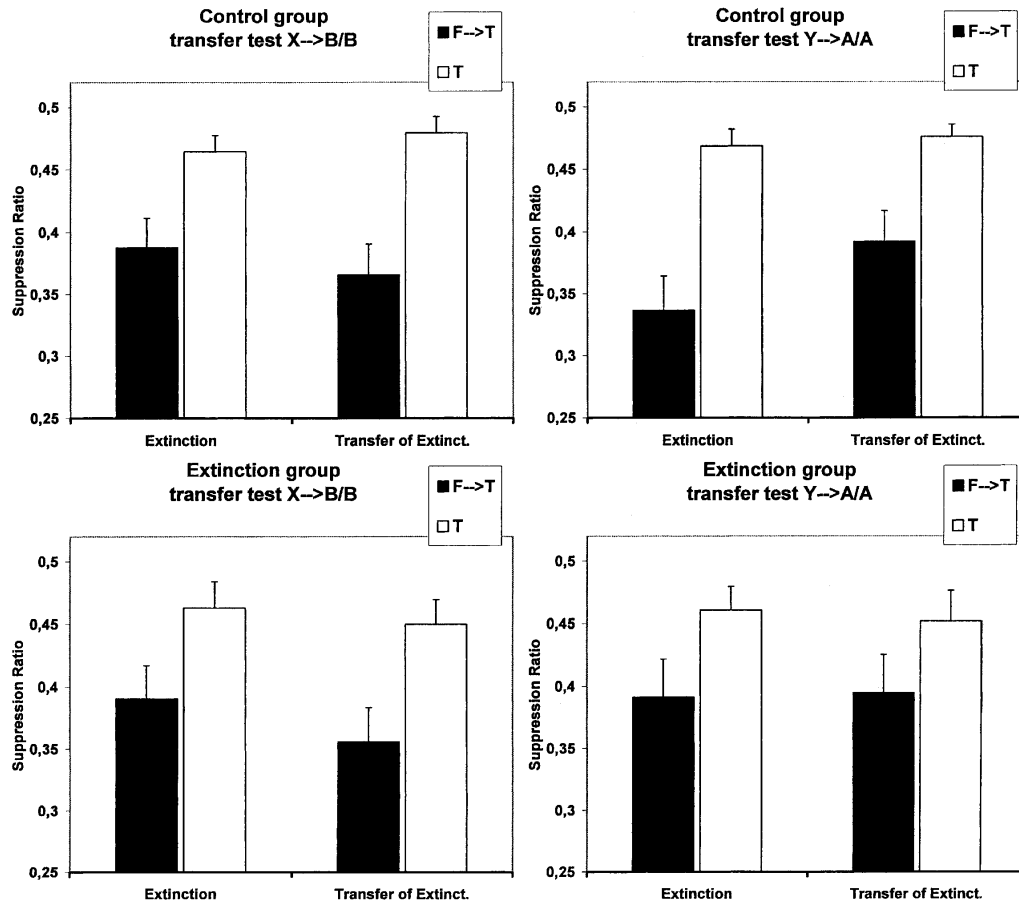


Figure 5. (Experiment 2) Mean suppression ratios at the tests of Extinction and Transfer of Extinction for the targets alone (T) and for the targets preceded by the feature (F→T), separately for the X→B/B (left side) and for the Y→A/A (right side) discriminations; the top half of the figure depicts the results of the Control group, the bottom half the results of the Extinction group. Lower values represent stronger conditioned responding. Brackets represent standard errors of the means.

There was no evidence for a change in differential responding across the two test contexts, $F_{(1,23)} < 1$, $MSE = 0.005$. Less surprisingly, we also observed differential responding to A preceded by Y versus to A alone both at the Extinction test, $F_{(1,23)} = 8.06$, $MSE = 0.007$, and at the Transfer of Extinction test, $F_{(1,23)} = 5.60$, $MSE = 0.007$. There was no sign of a change in differential responding across the two test moments, $F_{(1,23)} < 1$, $MSE = 0.004$.

Discussion

Participants in both conditions (Control/Extinction) readily mastered the two conditional discriminations (X→A+/A− and Y→B+/B−), and they again did so in a way indicative of occasion setting (strong suppression to target preceded by feature, combined with little suppression to either feature or target alone at the Acquisition test).

Next, in the Extinction group discriminative X→A/A and Y→B/B responding remained fully intact when assessed at the Transfer of Acquisition test in the new context *b*, whereas in the Control group feature Y's acquired facilitatory ability was similarly unaffected by the context switch and feature X showed some minor loss of facilitatory power. This pattern of results confirms the conclusion of Experiment 1 that positive modulation transfers well to a context different from the original acquisition context, and again makes it unlikely that the eventually observed *aba*-renewal would be due to incomplete extinction of the original excitatory/facilitatory associations and the residual associative strength showing up again in context *a*.

As in Experiment 1, in the Extinction group the nonreinforced feature–target presentations (X→A−) were clearly effective in behaviorally extinguishing the feature's (X) positive occasion-setting powers. Importantly, this loss of facilitation by feature X was observed while at the same time feature Y, which was given continued facilitatory Y→B+/B− training, was shown to fully retain its facilitatory power. Also, different from what was observed in Experiment 1 but in line with most other human and animal findings, in the Control group the nonreinforced feature-alone presentations (X−) did not affect at all the feature's positive occasion-setting powers. This set of results precludes an explanation of the loss of X's modulatory power in the Extinction group in terms of some general nonassociative mechanism, and supports Rescorla's (1986) conclusion obtained with pigeons that it is the conjunction of nonreinforcement and an excitatory target (X→A− treatment) that is critical to reduce a feature's positive occasion-setting properties. At the same time, these findings on the conditions of extinction of modulation attest to the plausibility of our hypothesis that the unexpected loss of modulatory ability of the control feature Y in Experiment 1 may, indeed, have been due to an acquired functional equivalence with feature X, which was given the causally effective X→A− extinction treatment.

As was observed in the *aba* condition of Experiment 1, in the Extinction group the modulatory ability of the extinguished feature X was again partially restored when assessed in a context *a* different from the extinction context *b* (Transfer of Extinction

test). This replication of *aba*-renewal of modulation confirms our previous conclusion that extinction of facilitation probably reflects contextually controlled suppression, rather than permanent erasure of the feature's modulatory ability. Importantly, this replication of renewal of Pavlovian modulation was obtained here under conditions that, contrary to Experiment 1, do allow for a straightforward interpretation of the loss of facilitatory power in terms of genuine associative extinction.

Finally, the tests of the features' effect on the targets that were different from those with which they had been trained originally ($X \rightarrow A$ and $Y \rightarrow B$ trials) delivered in part the predicted, in part an unexpected finding. As expected, in the Control condition (Extinction and Transfer of Extinction tests) the nonextinguished features X and Y facilitated responding to the new targets B and A , respectively. Hence, in line with what was observed in Experiment 1 with respect to features' renewed modulatory power, nonextinguished positive modulators also facilitated responding to targets that had been subject to modulation by other feature stimuli. Also in line with an anticipated outcome was the observation in the Extinction group that feature Y facilitated responding to target A both at the Extinction test in context b and at the Transfer of Extinction test in context a . This suggests that the nonreinforced $X \rightarrow A$ presentations had affected feature X 's modulatory associative link while—at least partially—preserving the target A -US associations, such that the nonextinguished feature Y was still able to facilitate responding to target A . The unexpected observation in the Extinction group was that the extinguished feature X , which no longer facilitated responding to its own target A , did reliably increase suppression to the transfer target B at the Extinction test in context b . This suggests that the effect of the $X \rightarrow A$ extinction trials may demonstrate some degree of target-specificity. This observation is not in line with the results of Experiment 1, however, where it was found that features that had lost the ability to modulate responding to the target they had been trained with, also failed to facilitate responding to transfer targets (*aba* group, Extinction test, $X \rightarrow B/B$ and $Y \rightarrow A/A$ trials).

Materials and Methods

Participants

A total of 48 undergraduate psychology students (35 females, 13 males, age 17–22) participated as a partial fulfillment of course requirements. None of them had any previous experience with the Martians preparation, and they were all uninformed as to the purpose of the experiment. An equal number of participants ($n = 24$) was assigned randomly to the Extinction and to Control groups. Within each group, for half of the participants, features X and Y were auditory stimuli and targets A and B were visual stimuli, whereas the reverse was true for the other half. Participants were tested individually, and test duration was ~45 min.

Apparatus, software, stimuli, and procedure

Experiment 2 made use of the same software, critical stimuli, context manipulations, and general procedure as Experiment 1. Hence pretraining, the US-only phase, $X \rightarrow A+/A-$ and $Y \rightarrow B+/B-$ conditional discrimination training in context a , the Acquisition test in context a , and also the Transfer of Acquisition test in context b were identical to those of the *aba* group of Experiment 1, and were the same for the Extinction and the Control conditions.

In the Extinction phase, participants in the Extinction condition were exposed in context b to 20 nonreinforced $X \rightarrow A-$ trials, to eight reinforced $Y \rightarrow B+$, and to eight nonreinforced $B-$ trials, the $X \rightarrow A-$ treatment aiming at extinguishing feature X 's facilitatory power, the $Y \rightarrow B+/B-$ treatment aiming at sustaining Y 's facilitatory power. The 20 $X \rightarrow A-$, eight $Y \rightarrow B+$, and eight $B-$ trials were subdivided into four equivalent sequential sub-blocks.

Each sub-block contained four $X \rightarrow A-$ trials, one $Y \rightarrow B+$ trial, and one $B-$ trial, followed by a single presentation each of $X \rightarrow A-$, $Y \rightarrow B-$, and $B-$ assessment trials. Semirandomization of trial types was accomplished in the same way as during the previous phases of the experiment, and the ITIs and stimulus presentation times were also identical to those used in the previous phases. Participants in the Control condition were treated in the same way as those in the Extinction group, the only crucial difference being that they received feature-alone ($X-$) trials instead of nonreinforced feature→target extinction trials ($X \rightarrow A-$) in context b . The $X-$ trials were intermixed with the $Y \rightarrow B+/B-$ trials in the same way as they were intermixed with the $X \rightarrow A-$ trials in the Extinction group. Both the $X-$ treatment and the continued $Y \rightarrow B+/B-$ training aimed at leaving X 's and Y 's facilitatory powers intact.

The Extinction test in context b and the final Transfer of Extinction test in context a followed the fourth extinction block without any interruption, and were identical to those of the *aba* group of Experiment 1.

Discussion

A first conclusion that follows from the present experiments pertains to the conditions that promote extinction of modulation resulting from sequential FP training. At first sight, the results of Experiment 1 seemed to threaten Rescorla's (1986) conclusion that it is the contingent pairing of the feature with a nonreinforced excitatory target stimulus that is necessary for extinction: also simple nonreinforced presentations of the feature alone ($Y-$ trials) were sufficient to remove the feature's facilitatory power. Our hypothesis was that this effect may have been due to a particular design characteristic of the experiment, namely, the within-subject manipulation of whether a feature was given $X \rightarrow A-$ extinction or $Y-$ control treatment after concurrent training on two similar sequential FP discriminations ($X \rightarrow A+/A-$ and $Y \rightarrow B+/B-$). The plausibility of the hypothesis that the unexpected loss of feature Y 's modulatory ability may have been due to an acquired functional equivalence between features X and Y was examined in Experiment 2, in which the same nonreinforced feature→target ($X \rightarrow A-$) versus feature-alone ($X-$) treatment was manipulated on a between-subjects base. The results of this study clearly showed that under those conditions, it was only the $X \rightarrow A-$ and not the $X-$ treatment that resulted in a loss of modulatory power of the feature. Taken together, these results on the conditions of extinction of modulation confirm in humans what has been found in a nonhuman population, and add to Rescorla's (1986) conclusion that the effect of nonreinforced feature→target presentations may spread to features that have a similar acquisition history. Finally, the fact that extinction of facilitation requires—at least in principle—that the feature is nonreinforced in the presence of an excitatory target, suggests that for an extinction treatment to be effective, it must contain information that is relevant and contradicts what has previously been learned. If what a positive occasion setter indicates is that another stimulus (the target) will soon be followed by the US, what is needed to contradict this is not the experience that the feature alone is not followed by reinforcement (which is irrelevant to and does not contract the information the feature carries), but the experience that the feature no longer indicates that the target stimulus will soon be followed by the US.

A second conclusion that follows from Experiments 1 and 2 is that extinction of Pavlovian modulation, just like extinction of simple Pavlovian excitation, definitely does not involve simple (complete) erasure or forgetting of what was previously learned. The fact that in both experiments a simple context switch was sufficient to (partially) restore extinguished facilitation rather suggests that, owing to $X \rightarrow A-$ extinction treatment, the associative structure underlying the feature's facilitatory power re-

mains (at least partially) intact, and is supplemented with new, contextually controlled corrective associations that counteract the effect and the (behavioral) expression of the original associative structure. In the present experiments, the context-control was shown using an *aba* renewal paradigm. Future research will have to determine whether it is critical for renewal of modulation to occur that the extinguished stimuli are retested in the acquisition context, or whether the phenomenon also shows up when tested in a third new context (*abc*-renewal), or maybe even when acquisition and extinction take place in the same context and the stimuli are then for the first time presented in a new context (*aab*-renewal). To the extent that the findings on the conditions of extinction and renewal of Pavlovian excitation can be generalized to Pavlovian modulation in humans, recent animal work at least suggests that each of these context manipulations may produce renewal of facilitation, but that the *aba*-renewal demonstrated here will probably produce the largest effect (e.g., Thomas et al. 2003).

Third, even though the present studies were not designed to unravel the nature of the representational structure underlying extinction of modulation, some preliminary suggestions can be made concerning the locus of intervention of the corrective associations that counteract the expression of the original associative structure. Given the observation (Experiment 2, Extinction group, Extinction test) that the nonextinguished feature (Y) was still able to facilitate responding to the target (A) that had been involved in the $X \rightarrow A$ – extinction treatment, it is likely that extinction of modulation involves a mechanism that acts directly on the modulatory associative link. In terms of the model depicted in Figure 1A, this would imply that the direct excitatory and the gated inhibitory target–US associations remain at least partially intact, and that the inhibitory influence of the feature on the inhibitory target–US association is suspended as a consequence of $X \rightarrow A$ – extinction treatment. Following this assumption, a nonextinguished feature (such as Y) could still inhibit the inhibitory target–US association, and thus produce activation of the US and elicit conditioned responding. Moreover, given the observed contextual control on the extinguished feature's facilitatory ability, a type of modulated modulatory structure seems required, whereby the activity of the new link that suspends the influence of the original modulatory associative link in its turn can be controlled by contextual cues. Finally, the unexpected observation (Experiment 2, Extinction group, Extinction test) that the extinguished feature X, which no longer facilitated responding to its own target A, was still able to increase suppression to the transfer target B suggests an additional mechanism that may be involved in extinction of modulation: strengthening of the original inhibitory target A–US association. Namely, to the extent that the loss of conditioned responding to target A preceded by feature X would be the result of the joint contribution of a partial suppression of the feature's modulatory influence combined with A–US inhibition having grown stronger, feature X might still be able to promote responding to a transfer target B. Said differently, if the suppression of the feature's modulatory power is incomplete, feature X might still be able to act as a substitute for feature Y and thus increase suppression to the transfer target B, even when tested in the context of extinction. However, because this phenomenon was observed in Experiment 2 but not at all in Experiment 1, future research will have to confirm or to reject the plausibility of the latter mechanism.

As a fourth conclusion, we want to stress that the present findings add to the growing body of evidence that, even though modulators and excitors/inhibitors fundamentally differ in the content of the information they carry, from this it does not necessarily follow that the associative processes that are operative in the establishment of modulation should follow laws fundamen-

tally different from those applying to the acquisition of simple Pavlovian excitation/inhibition. On the contrary, Miller and Oberling (1998) and Bonardi (1998) even argued that for an association between a feature and a representation of the complex event “target–US,” similar if not identical laws apply as for associations between a CS and a representation of the simple event “US.” For example, just as simple CSs are subject to cue competition, there also exists competition (such as blocking and overshadowing) between potential occasion setters (Bonardi 1998; Miller and Oberling 1998). As a matter of fact, Miller and Oberling (1998) argue for many other analogies between Pavlovian excitation and occasion setting, including evidence of temporal encoding, latent inhibition, learned irrelevance, modulation by higher-order stimuli, summation, and time ratio effects. To this list we can now add extinction, and contextually controlled renewal of extinguished responding.

Fifth and finally, at a more general level, we want to stress the striking similarities between these human data on acquisition and extinction of positive modulation and what is typically found in nonhuman animal preparations (see also Baeyens et al. 2001). This may come as no surprise to those working with the assumption of cross-species generality of the basic learning processes. However, parsimonious and fruitful as the assumption of cross-species generality may be in principle, the replication of several key findings from the animal laboratory has, as a matter of fact, not always been obvious in human conditioning preparations (e.g., latent inhibition [Lubow and Gewirtz 1995], blocking [Arcediano et al. 1997], sensitivity to schedules of reinforcement [Hayes 1989]). From this perspective, the present observation that humans apparently do behave as predicted by current animal models of occasion setting with respect to FP conditional discrimination learning, shows that it still may be worthwhile to temporarily bracket the fact that people think and verbalize explicit rules, and to see how far one can get in explaining (also more complex) human learning with a purely associative account. This is a reassuring thought for associative learning theorists, but probably no less for those behavioral neuro-scientists who use animal conditioning preparations and also hope to contribute to an understanding of learning in humans.

Acknowledgments

Debora Vansteenwegen and Tom Beckers are postdoctoral researchers for the Fund of Scientific Research (Flanders).

References

- Arcediano, F., Ortega, N., and Matute, H. 1996. A behavioural preparation for the study of human Pavlovian conditioning. *Q. J. Exp. Psych.* **49B**: 270–283.
- Arcediano, F., Matute, H., and Miller, R.R. 1997. Blocking of Pavlovian conditioning in humans. *Learn. Motiv.* **28**: 188–199.
- Baeyens, F. and Clarysse, J. 1998. *Martians for Windows™95 computer program*. University of Leuven, Leuven, Belgium.
- Baeyens, F., Vansteenwegen, D., Hermans, D., Vervliet, B., and Eelen, P. 2001. Sequential and simultaneous Feature Positive discriminations: Occasion setting and configural learning in human Pavlovian conditioning. *J. Exp. Psychol. Anim. Behav. Process.* **27**: 279–295.
- Baeyens, F., Vansteenwegen, D., Beckers, T., Hermans, D., and Vervliet, B. 2005. No effects of a context switch following sequential Feature-Negative nor Feature-Positive discrimination training: Evidence for contextual transfer of target inhibition. In *Progress in learning research* (ed. F. Columbus). Nova Science Publishers, Inc., Hauppauge, NY (in press).
- Bonardi, C. 1998. Conditional learning: An associative analysis. In *Occasion setting: Associative learning and cognition in animals* (eds. N.A. Schmajuk and P.C. Holland), pp. 37–67. APA, Washington, DC.
- Bouton, M.E. 1993. Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychol. Bull.* **114**: 80–99.
- . 2000. A learning theory perspective on lapse, relapse, and the maintenance of behavior change. *Health Psychol.* **19**: 57–63.

- . 2002. Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biol. Psychiatry* **51**: 976–986.
- Bouton, M.E. and King, D.A. 1983. Contextual control of the extinction of conditioned fear: Tests for the associative value of the context. *J. Exp. Psychol. Anim. Behav. Process.* **9**: 248–265.
- Bouton, M.E. and Nelson, J.B. 1998. Mechanisms of Feature-Positive and Feature-Negative discrimination learning in an appetitive conditioning paradigm. In *Occasion setting: Associative learning and cognition in animals* (eds. N.A. Schmajuk and P.C. Holland), pp. 69–112. APA, Washington, DC.
- Bouton, M.E. and Peck, C.A. 1989. Context effects on conditioning, extinction, and reinstatement in an appetitive conditioning preparation. *Anim. Learn. Behav.* **17**: 188–198.
- Bouton, M.E. and Swartzentruber, D. 1986. Analysis of the associative and occasion-setting properties of contexts participating in a Pavlovian discrimination. *J. Exp. Psychol. Anim. Behav. Process.* **12**: 333–350.
- . 1989. Slow reacquisition following extinction: Context, encoding and retrieval mechanisms. *J. Exp. Psychol. Anim. Behav. Process.* **15**: 43–53.
- Bouton, M.E., Kenney, F.A., and Rosengard, C. 1990. State-dependent fear extinction with two benzodiazepine tranquilizers. *Behav. Neurosci.* **104**: 44–55.
- Cunningham, C.L. 1979. Alcohol as a cue for extinction: State dependency produced by conditioned inhibition. *Anim. Learn. Behav.* **7**: 45–52.
- Estes, W.K. 1976. Structural aspects of associative models of memory. In *The structure of human memory* (ed. C.N. Cofer), pp. 31–53. Freeman, New York.
- Hall, G. 1996. Learning about associatively activated representations: Implications for acquired equivalence and perceptual learning. *Anim. Learn. Behav.* **24**: 233–255.
- Hayes, S.C. 1989. *Rule-governed behavior: Cognition, contingencies, and instructional control*. Plenum Press, New York.
- Holland, P.C. 1983. "Occasion-setting" in Pavlovian Feature-Positive discriminations. In *Quantitative analyses of behavior: Vol. IV. Discrimination processes* (eds. M.L. Commons et al.), pp. 183–206. Ballinger, Cambridge, MA.
- . 1986a. Temporal determinants of occasion setting in Feature Positive discriminations. *Anim. Learn. Behav.* **14**: 111–120.
- . 1986b. Transfer after serial Feature Positive discrimination training. *Learn. Motiv.* **17**: 243–268.
- . 1989a. Feature extinction enhances transfer of occasion setting. *Anim. Learn. Behav.* **17**: 269–279.
- . 1989b. Transfer of negative occasion setting and conditioned inhibition across conditioned and unconditioned stimuli. *J. Exp. Psychol. Anim. Behav. Process.* **15**: 311–328.
- . 1991. Acquisition and transfer of occasion setting in operant Feature Positive and Feature Negative discriminations. *Learn. Motiv.* **22**: 366–387.
- . 1992. Occasion setting in Pavlovian conditioning. *Psychol. Learn. Motiv.* **28**: 69–125.
- Honey, R.C. and Hall, G. 1989. Acquired equivalence and distinctiveness of cues. *J. Exp. Psychol. Anim. Behav. Process.* **15**: 338–346.
- Kirk, R.E. 1995. *Experimental design: Procedures for the behavioral sciences*. Brooks/Cole, Pacific Grove, CA.
- Lamarre, J. and Holland, P.C. 1987. Transfer of inhibition after serial Feature Negative discrimination training. *Learn. Motiv.* **18**: 319–342.
- Lamoureux, J.A., Buhusi, C.V., and Schmajuk, N.A. 1998. A real-time theory of Pavlovian conditioning: Simple stimuli and occasion setters. In *Occasion setting: Associative learning and cognition in animals* (eds. N.A. Schmajuk and P.C. Holland), pp. 383–424. APA, Washington, DC.
- Lubow, R.E. and Gewirtz, J.C. 1995. Latent inhibition in humans: Data, theory, and implications for schizophrenia. *Psychol. Bull.* **117**: 87–103.
- Matute, H. and Pineno, O. 1998. Stimulus competition in the absence of compound conditioning. *Anim. Learn. Behav.* **26**: 3–14.
- Miller, R.R. and Oberling, P. 1998. Analogies between occasion setting and Pavlovian conditioning. In *Occasion setting: Associative learning and cognition in animals* (eds. N.A. Schmajuk and P.C. Holland), pp. 3–35. APA, Washington, DC.
- Nakajima, S. 1994. Contextual control of Pavlovian Feature-Positive and Feature-Negative discriminations. *Anim. Learn. Behav.* **22**: 34–46.
- . 1998. Pavlovian feature-ambiguous discrimination. In *Occasion setting: Associative learning and cognition in animals* (eds. N.A. Schmajuk and P.C. Holland), pp. 147–165. APA, Washington, DC.
- Pavlov, I.P. 1927. *Conditioned reflexes*. Oxford University Press, London.
- Pearce, J.M. 1987. A model for stimulus generalization in Pavlovian conditioning. *Psychol. Rev.* **94**: 61–75.
- Rescorla, R.A. 1985. Conditioned inhibition and facilitation. In *Information processing in animals: Conditioned inhibition* (eds. R.R. Miller and N.E. Spear), pp. 299–326. Erlbaum, Hillsdale, NJ.
- . 1986. Extinction of facilitation. *J. Exp. Psychol. Anim. Behav. Process.* **12**: 16–24.
- . 1987. Facilitation and inhibition. *J. Exp. Psychol. Anim. Behav. Process.* **13**: 250–259.
- . 1991. Transfer of inhibition and facilitation mediated by the original target stimulus. *Anim. Learn. Behav.* **19**: 65–70.
- Schmajuk, N.A., Lamoureux, J.A., and Holland, P.C. 1998. Occasion setting: A neural network approach. *Psychol. Rev.* **105**: 3–32.
- Swartzentruber, D. 1995. Modulatory mechanisms in Pavlovian conditioning. *Anim. Learn. Behav.* **23**: 123–143.
- Swartzentruber, D. and Rescorla, R.A. 1994. Modulation of trained and extinguished stimuli by facilitators and inhibitors. *Anim. Learn. Behav.* **22**: 309–316.
- Thomas, B.L., Larsen, N., and Ayres, J.J.B. 2003. Role of context similarity in ABA, ABC, and AAB renewal paradigms: Implications for theories of renewal and for treating human phobias. *Learn. Motiv.* **34**: 410–436.
- Wagner, A.R. 1981. SOP: A model of automatic memory processing in animal behavior. In *Information processing in animals: Memory mechanisms* (eds. N.E. Spear and R.R. Miller), pp. 5–47. Erlbaum, Hillsdale, NJ.

Received April 27, 2004; accepted in revised form February 8, 2005.